

# **EXHIBIT F**

**UNITED STATES DISTRICT COURT  
DISTRICT OF NEW JERSEY**

IN RE: VALSARTAN,  
LOSARTAN, AND IRBESARTAN  
PRODUCTS LIABILITY  
LITIGATION

HON. ROBERT B. KUGLER  
HON. JOEL SCHNEIDER

Civil No. 19-2875 (RBK/JS)

---

**DECLARATION OF MAURA R. GROSSMAN IN FURTHER SUPPORT OF  
THE TEVA DEFENDANTS' MOTION TO ENFORCE ESI PROTOCOL AND  
SUPPLEMENTAL LETTER BRIEF**

Pursuant to 28 U.S.C. § 1746, I, Maura R. Grossman, declare:

1. I am a Research (Full) Professor of Computer Science in the David R. Cheriton School of Computer Science at the University of Waterloo and an Adjunct Professor of Law at Osgoode Hall Law School at York University, both in Ontario, Canada. I also am an eDiscovery attorney and consultant in Buffalo, New York. Prior to establishing my own law and consulting firm in June 2016, for 17 years I was a litigator at Wachtell, Lipton, Rosen & Katz in New York City, where I oversaw the firm's eDiscovery efforts. I have practiced exclusively in the area of eDiscovery for the past 14 years.

2. I have been recognized as one of the foremost authorities on technology-assisted review ("TAR"). In *Rio Tinto v. Vale*, No. 14 Civ. 3042 (RMB) (AJP), 2015 WL 4367250 (S.D.N.Y. July 15, 2015), Magistrate Judge Andrew J. Peck appointed me as a special master to resolve disputes that arose in connection with the use of TAR, and in doing so, referred to me as "one of the most knowledgeable lawyers (if not *the* most knowledgeable lawyer) about TAR. . . ." *Id.* at \*1 (emphasis in original).

3. In 2010, my colleague at the University of Waterloo, Professor Gordon V. Cormack, and I invented continuous active learning (“CAL”)—often referred to as “TAR 2.0”—for technology-assisted review. I hold ten (10) patents related to that technology and own the registered trademarks for “continuous active learning” and “CAL.” Nonetheless, I derive no benefit, financial or otherwise, from any party’s use of CAL.

4. Since 2010, I have been involved in well over one-hundred (100) TAR reviews, the vast majority of which have involved CAL. I have over ninety (90) publications on the subject of eDiscovery and more than twenty (20) peer-reviewed publications in the scientific literature relating to CAL. My research on TAR with Professor Cormack has been cited approvingly in many federal and state court cases, including *Da Silva Moore v. Publicis Groupe*, 287 F.R.D. 182, 190 (S.D.N.Y. 2012), the first case to recognize and approve the use of TAR in civil litigation. Our study, *Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review*, published in volume 17 of the *Richmond Journal of Law and Technology* (2011)<sup>1</sup> has been widely cited in case law, both in the U.S. and abroad. My TAR research is funded by nine grants and fellowships, including those from the Ontario Research Fund, the Canadian Foundation for Innovation, and the National Sciences and Engineering Research Council of Canada.

5. I have conducted TAR projects and served as a TAR expert both for plaintiffs and defendants, and have been appointed as a neutral, expert to the court, or special master in a dozen high-profile federal and state court cases, many of which have involved resolving disputes related to search methodologies, including TAR.

6. For the sake of brevity, I refer the Court to my *curriculum vitae*, attached hereto as **Exhibit A**, setting forth my additional qualifications to serve as a TAR expert in this matter, particularly as relates to the application and validation of CAL.

---

<sup>1</sup> For the sake of brevity, the publications referenced in this declaration are not attached as exhibits. However, to the extent the Court would like to review any of the referenced articles, the Teva Defendants will promptly submit them.

7. I have been retained by Greenberg Traurig in the above-captioned litigation to assist the Teva Defendants on matters relating to search and review methodologies involving CAL, in particular, to develop a validation protocol and to oversee the Teva Defendants' review process to ensure that the Teva Defendants' productions are reasonable, made in good faith, and consistent with their obligations under Fed. R. Civ. P. 26(g). I have negotiated and/or developed validation protocols in the past, including but not limited to those in *Rio Tinto v. Vale*, Case No. 1:14-cv-03042-RMB-AJP (S.D.N.Y. Sept. 8, 2015) (Stipulation and Order Re: Revised Validation and Audit Protocols for the Use of Predictive Coding in Discovery [Dkt. 338]), and *In re Broiler Chicken Antitrust Litigation*, Case No. 1:16-cv-08637 (N.D. Ill. Jan. 3, 2018) (Order Regarding Search Methodology for Electronically Stored Information [Dkt. 586]).

8. I submit this declaration on behalf of the Teva Defendants in Support of their Motion to Enforce ESI Protocol [Dkt. 516] and supplemental Letter Brief of July 24, 2020 [Dkt. 527]. I am familiar with the issues raised in the Parties' current dispute involving the use of CAL. I write to address certain claims made in Plaintiffs' Letter Brief of July 24, 2020 [Dkt. 526] and in the accompanying Declaration of Jonathan Jaffe (Ex. A thereto), which I believe to be unfounded and materially incorrect in numerous ways. I have personal knowledge of the matters set forth herein, and if called upon and sworn as a witness, I could and would competently testify thereto.

9. Professor Cormack and I first coined the term "technology-assisted review" or "TAR" in our 2011 Richmond Journal paper (referenced in ¶ 4 above). We subsequently defined the term more formally in *The Grossman-Cormack Glossary of Technology-Assisted Review*, published in volume 7 of Federal Courts Law Review (2013). At its core, TAR is a process for ranking—from most to least likely to be responsive—or for classifying—as responsive or non-responsive—a document collection, using computer software that learns to distinguish between responsive and non-responsive documents based on coding decisions made by one or more knowledgeable reviewers on a subset of the document collection. The software then applies what it has learned to the remaining documents in the collection. While TAR tools differ in their specific

mechanisms and workflows, all TAR tools either rank or classify documents based on their likelihood of being responsive.

10. The TAR training process typically begins with the review by one or more knowledgeable attorneys of a set of documents that includes diverse examples of both responsive and non-responsive documents drawn from the universe of potentially responsive documents. As training continues, knowledgeable attorneys review additional batches of documents for responsiveness, on an iterative basis, and as the attorneys review more documents, the TAR tool continually improves in its ability to distinguish between responsive and non-responsive documents. This learning is applied to the documents in the collection that have not yet been reviewed by an attorney. Training continues until the TAR tool can no longer find responsive documents, and the few responsive documents that remain in the collection are so sparse that the effort to find them is no longer proportionate. With CAL, as opposed to TAR 1.0, there is no distinction between “training” and “review”; as documents are reviewed and coded as responsive or nonresponsive by attorneys, the coding is continually fed back into the CAL system to improve its ability to distinguish between responsive and non-responsive documents. Therefore Mr. Jaffe’s statements (Decl. at 16 ¶¶ 9, 10) about producing documents exceeding a certain threshold score without review is a misconception, which indicates a lack of understanding of the CAL process, because with CAL, the scores change after batches are coded, until the very last document is coded. There is no way to do what Mr. Jaffe is demanding, as it applies only to TAR 1.0.

11. For the Court’s benefit, below is a table showing the primary differences between TAR 1.0 and TAR 2.0 (CAL):

TAR 1.0	TAR 2.0 (CAL)
<p><b>1. Separate Training and Review Stages.</b> Training and review are completely different phases; TAR 1.0 does not provide for further training once stabilization has been reached and review begins.</p>	<p><b>1. Training = Review and Review = Training.</b> Continuous Active Learning permits the TAR algorithm to keep learning and improving over the course of the review.</p>
<p><b>2. Trains Against a Control Set.</b> Limits the ability to handle rolling uploads; assumes all documents will be collected before training begins; stabilization is measured against performance on a Control Set.</p>	<p><b>2. Re-Ranks All of the Documents Repeatedly.</b> Allows for rolling uploads and course correction; does not use a Control Set, but rather, monitors precision as the review process proceeds and validates the production at the end of the process.</p>
<p><b>3. Expensive Subject Matter Expert (“SME”) Handles Training.</b> Review team judgments are not used to train the system, only for the review phase.</p>	<p><b>3. Review Teams Handle Training as They Review.</b> Reviewers work alongside SMEs for maximum effectiveness; SMEs focus on the initial training rounds and QCing review team judgments, <i>e.g.</i>, conflicts between the TAR classification score and the reviewer coding decisions.</p>
<p><b>4. Typically Uses Random Seeds to Train the System,</b> rather than key documents found by the legal team, so training is slow and expensive.</p>	<p><b>4. Uses Diverse Judgmental Seeds to Train the system,</b> quickly identifying highly relevant documents; continues until no more relevant documents can be found.</p>
<p><b>5. Does Not Work Well with Low Richness/Prevalence Collections;</b> impractical for small cases or cases with few relevant documents because of up-front training requirement.</p>	<p><b>5. Works Well with Low Richness/Prevalence Collections;</b> useful for any size case, regardless of the number of relevant documents, because of the flexible workflow and low initial overhead.</p>

12. After the review process is complete, TAR efforts are typically evaluated using two metrics drawn from the science of information retrieval. One is referred to as *recall*, which is a measure of completeness, reflected by the proportion (*i.e.*, percent) of responsive documents that

have been found through a search or review process, out of all possible responsive documents in the collection. The other is referred to as *precision*, which is a measure of accuracy, or the proportion (*i.e.*, percent) of the documents identified by a search or review process that are actually responsive. High recall suggests that substantially all responsive documents have been found; high precision suggests that primarily responsive documents have been found. Low precision, on the other hand, suggests that there are many non-responsive documents in the production set. Most TAR productions seek to achieve high recall, with a reasonable level of precision. In order to determine whether a TAR process is adequate, validation (through statistical sampling) is typically performed at the conclusion of the TAR effort to confirm that there are not an excessive number of responsive documents in the set labelled as non-responsive by the TAR tool (*i.e.*, low recall), and that there are not an excessive number of non-responsive documents in the production (*i.e.*, low precision). Courts have routinely accepted statistical sampling as a valid means of measurement when the population of interest (in this case, in excess of eight million documents) is too large to count.

13. The Teva Defendants are using a CAL method, not TAR 1.0, to review custodial documents in this matter. The Rand Corporation has estimated that attorney review time accounts for \$0.73 out of every dollar spent on electronic discovery. On average, the cost of a contract attorney to review a single document is about \$1.00; law firm associates are considerably more expensive (*i.e.*, as much as \$4.00 to \$5.00 per document). Human reviewers typically can review no more than 50 documents an hour. Teva Defendants' vendor, Consilio, has attested that their current review rate is 43 documents an hour. They have estimated that the cost to manually review the approximately 3.7 million documents that will result after the negotiated keywords have been applied to the entire document collection will exceed \$6 million and will take 11 or more months to complete. These estimates are consistent with my experience. My research has shown that CAL can be as much as *50 times more efficient than manual review, and also more effective*. See Richmond Journal study cited in ¶ 4 above; *see also* Gordon V. Cormack & Maura R. Grossman,

*Navigating Imprecision in Relevance Assessments on the Road to Total Recall: Roger and Me*, in Proc. of the 40th Int'l ACM SIGIR Conference on Rsch. & Dev. in Info. Retrieval 5 (2017).

14. As mentioned above, Plaintiffs and their consultant appear to have a fundamental misconception about how CAL works and many of their statements relate to TAR 1.0, not to CAL. Brainspace's implementation of CAL—continuous multimodal learning (“CMML”)—the TAR tool that the Teva Defendants have proposed to use in this matter, employs an algorithm called logistic regression. Logistic regression is a *discriminative* classifier, which means that it looks for distinct *features* that distinguish responsive from non-responsive documents. It is not, as Mr. Jaffe suggests, a *generative* classifier that seeks to find documents that “look like” the responsive training examples. In the current matter, words like “NDMA,” “nitrosodimethylamine,” “valsartan,” “hydrochlorothiazide,” and “contamination” are far more likely to occur in responsive documents, while words like “Copaxone,” “Azilect,” and “groundwater” are far more likely to occur in non-responsive documents. Words like “test” and “report” would not occur disproportionately in either responsive or non-responsive documents. In his contrived example (Decl. at 5), Mr. Jaffe posits that if reviewers have coded 1,000 test reports and determined that not a single test report is relevant, thereafter the system will exclude reports containing features like “NDMA” that indicate responsiveness. That is not how CAL works. CAL selects the most-likely responsive documents first for review, so that any test report containing NDMA and related terms would be selected for review before a test report not containing any of the words or features indicating responsiveness. As the review progresses, CAL might present non-NDMA test results and would learn to exclude them, not because they were test results, but because they contained terms like “Copaxone” or “groundwater,” and not terms like “NDMA.” In the end, features like “test” and “result” would have very little impact on the CAL classification score because they appear commonly in *both* responsive and non-responsive documents.

15. Mr. Jaffe's misunderstandings about the nature of CAL pervade his Declaration and lead to a flood of inaccuracies and a parade of horrors concerning what CMML is likely to do in this matter. Mr. Jaffe makes a variety of assertions that are lacking in empirical support and are



inconsistent not only with my experience in conducting over 100 TAR projects, but my scientific research as well. For example, Mr. Jaffe argues that highly relevant documents will get a low score because they look like other irrelevant documents. While that might be true if an algorithm were designed to look for “more like this,” CMML does not do so. It looks for features that discriminate between responsive and non-responsive documents. Rare documents that share features in common with other responsive documents will easily be found by CMML; therefore, Mr. Jaffe’s concern is unfounded.

16. While Mr. Jaffe is correct that the scores assigned to documents relate to the likelihood of relevance and not necessarily the degree of relevance, he neglects to state that the two are highly correlated and that a CAL process, unlike TAR 1.0, is designed to elevate the most likely to be relevant and therefore, the most highly relevant documents, before documents that are only marginally relevant. It is TAR 1.0 that focuses on drawing the boundary between marginally relevant and non-relevant documents. Thus, Mr. Jaffe’s statements (Decl. at 6) that (i) “the scoring of a document has no correlation to how relevant a document is”; (ii) that Defendants “representation that their CMML/CAL process will automatically prioritize the production of **the most relevant** documents does not stand the weight of even a cursory examination,” (Decl. at 7 (emphasis in original)); and (iii) that “there is no reasonable basis to assert that the prioritized documents are **the most relevant**,” (*id.* (emphasis in original)) are simply wrong. Moreover, Mr. Jaffe fails to explain how manual reviewers conducting a linear review of millions of documents could find the most relevant documents to be reviewed first. Mr. Jaffe’s analogy to finding needles in a haystack also fails. He claims that if one is looking for needles in a haystack, the best way to do so is for humans to grab handfuls of hay and to search for the needles. I disagree. The best way to find the needles would be to use a metal detector, which is the equivalent of what CAL does. It trains the system to identify the needles, both sharp and dull, by looking for *metal objects*, a *feature* of needles, rather than by looking for things that *look like* needles. The Teva Defendants’ vendor has confirmed they are using diverse sets of responsive and non-responsive exemplars (*i.e.*, both sharp and dull needles, as well as different-shaped hay) to train the CMML system. That

process has also involved the use of the keywords previously negotiated by the parties to identify such documents. If Plaintiffs want to propose a reasonable set of training documents for the Teva Defendants' consideration for these purposes, my understanding is that the Teva Defendants are happy to consider that request.

17. Mr. Jaffe spills much ink on the issue of "bias," without a shred of support that it negatively impacts CAL (*see, e.g.*, Decl. at 8-9, 13-14); he seems to misunderstand that the whole point of CAL is to "bias" the system to find responsive documents. Taking GPS as an example, it would make no sense to drive in a random direction; one "biases" the GPS system to drive in the direction one is interested in going. But even if one drove 1,000 miles in the wrong direction, the GPS system would still correct that error. My experience in conducting over 100 TAR matters is inconsistent with Mr. Jaffe's "bias" theory. It would take a very determined effort by the Teva Defendants to mislead a CAL system so as not to find responsive documents, whether highly or marginally so.

18. Mr. Jaffe is also wrong in the places he discusses the requirement that Defendants must use "multiple tags/vectors" to train a CAL system (Decl. at 12, 14-15 ¶ 1). I have already noted the importance of training the system on diverse exemplars representing the various aspects of relevance, but *I have never seen or engaged in a TAR process that required the training of a different model for each aspect of relevance*. CMML already includes a process to sample from different portions of the collection to identify training documents, which should be perfectly adequate to address Mr. Jaffe's concerns about diversity and density. Moreover, my research demonstrates that CAL is able to find all aspects of relevance, and that doing so using a single relevant/non-relevant distinction is equally effective and more efficient than the approach Mr. Jaffe advocates. *See* Gordon V. Cormack & Maura R. Grossman, *Multi-Faceted Recall of Continuous Active Learning for Technology-Assisted Review*, in Proc. of the 38th Int'l ACM SIGIR Conference on Rsch. & Dev. in Info. Retrieval 763 (2015). While it is correct that a CAL tool may find some topics before others, it will eventually find the vast majority of relevant documents across all aspects of relevance.

19. Mr. Jaffe's concern (Decl. at 6, 16) that Defendants will run TAR on documents that are unsuitable is a red herring and is ill founded. Mr. Jaffe is correct that TAR tools need text to function properly, but he is not correct to assume *ipse dixit* that TAR cannot work on hard-copy documents that have been OCR'd, spreadsheets, or documents with little text. Even a chromatograph is likely to be accompanied by a key or caption, or attached to a cover email, containing a relevant feature. I have used TAR on OCR'd documents, spreadsheets, documents with embedded charts and graphs, and short documents (including text messages and tweets) without a problem. My research also confirms that TAR can be used effectively on short documents and spreadsheets. *See id.* That said, whether a particular TAR tool is effective with particular file types is very much a tool-specific issue. The Teva Defendants' vendor has advised me that they have and will continue to remove document types that are not suitable for CMML, including but not limited to, images and certain spreadsheets, and that they will subject those to a separate manual-review workflow.

20. Mr. Jaffe's claims—(i) that “if reviewers see a consecutive series of low scored documents that they agree are irrelevant, they will begin to ‘trust’ the computer’s judgment. The weight of marking rarer documents as irrelevant has a much stronger adverse effect than the positive effect of marking a rare document relevant in a CAL review” (Decl. at 6), and (ii) that “[a]n irrelevant document in a sea of relevant documents is more likely to be marked as relevant. Conversely, and of greater concern as to bias against the Plaintiffs, relevant documents, even highly relevant, are less likely to be detected as relevant by a reviewer if they are presented in a sea of documents deemed irrelevant” (Decl. at 8-9)—are contrary to empirical evidence, which shows exactly the opposite effect. When reviewers see a string of relevant documents, they become more discriminating and tend to tag fewer documents as responsive, and when reviewers see a string of non-relevant documents they become less so, flagging more documents as responsive. *See Adam Roegiest & Gordon V. Cormack, Impact of Review-Set Selection on Human Assessment for Text Classification*, in Proc. of the 39th Int’l ACM SIGIR Conference on Rsch. & Dev. in Info. Retrieval 861 (2016). I have confirmed that the Teva Defendants' quality control

(“QC”) process will include the review of samples of documents classified by CMML with high scores that are marked non-responsive by humans, and samples of documents classified with low scores by CMML that are labeled as responsive by reviewers. This QC process should catch and correct for the most-likely coding errors.

21. Mr. Jaffe states (Decl. at 17), without providing a single reason, that the TAR protocol used in the *Broiler Chicken Antitrust Litigation*—one that I negotiated and ordered as the special master in that case—“was very limited and inadequate for a litigation of this size and complexity.” *Broilers* is a highly complex antitrust class action matter involving three separate classes, dozens of direct action plaintiffs, and more than 20 different defendants, with productions the same size if not larger than the ones anticipated here. A number of the defendants in *Broilers* used CAL methods for the purposes of conducting their document reviews. The parties and the Court agreed that the validation protocol in *Broilers* was state of the art and highly effective. I have negotiated other, similar TAR validation protocols, and contrary to Mr. Jaffe’s claim that devising a TAR validation protocol post production would be “infeasible,” or would be “unworkable at this point in the process” (Decl. at 17, 18.) in *Rio Tinto v. Vale*, the validation protocol was negotiated *after the review process was completed*. If Plaintiffs are willing to be reasonable and to cooperate, I see nothing that would prevent the parties from quickly negotiating an effective and proportionate validation protocol that would ensure that the Teva Defendants’ production is reasonable, made in good faith, and consistent with their obligations under Fed. R. Civ. P. 26(g). Nor should this negotiation cause any delay in the Teva Defendants’ rolling document productions or in meeting the Court’s deadlines.

22. In sum, many of the requirements Mr. Jaffe seeks to impose on the Teva Defendants in the section of his Declaration labeled “CAL Properly Done” (Decl. at 14-17) are either applicable solely to TAR 1.0, or are inconsistent with either standard or best practices in the use of CAL. Many are downright counterproductive and will lead to outcomes that are contrary to what Plaintiffs claim they want; others will inevitably delay and bog down the TAR process. For example, I have never seen or conducted a TAR review in which tags were weighted in the

manner that Mr. Jaffe describes (Decl. at 15 ¶ 2), nor do I see any benefit to doing so. In my experience, there is no material difference between a TAR model trained using a “responsive” tag versus a “hot” tag. Moreover, Mr. Jaffe provides no examples of any TAR protocols that have incorporated such a methodology. Because CAL systems are resistant to error, the Teva Defendants’ QC process will provide for timely course correction, and the Teva Defendants will employ a sound validation protocol at the conclusion of the TAR process, any need for joint training by the parties is obviated. While such joint training was employed in the TAR 1.0 context, when TAR was initially judicially approved in 2012, it has become far less common in the past few years, and with CAL, is unnecessary and counterproductive. Indeed, if Plaintiffs were to review the docket in the *Actos* case they cite in both their Letter Brief of July 24, 2020 [Dkt. 526] and their letter submitted to the Court on July 14, 2020 [Dkt. 513], they would learn that the joint-review process turned out to be unworkable and was ultimately abandoned.

23. Nonetheless, some of Mr. Jaffe’s requests could be accommodated by the Teva Defendants, and they have advised me that they are already doing or are willing to do some of them. For example, the Teva Defendants have informed me that they have already agreed to run CMML over the entire data set, rather than over only the documents matching the negotiated search terms. The Teva Defendants have also advised me that they are already implementing Plaintiffs’ request for prioritization through the selection of custodians. The Teva Defendants’ vendor has already been removing and will continue to remove file types that are unsuitable for CMML, and those documents will be addressed through a separate manual-review workflow. CMML already has a diversity training feature that includes documents that are dissimilar to the documents already trained on during the course of the review process. As Mr. Jaffe has requested, the documents that have been coded by reviewers are already, as a matter of course, fed back into the CMML system to support further training. As previously mentioned, the Teva Defendants’ review process will involve robust QC and validation mechanisms. There are others demands, however, that simply cannot be accommodated. For example, the Teva Defendants cannot produce documents above a certain threshold score without review because in a CAL process, by definition, the scores are

continually readjusting and all potentially responsive documents must be reviewed. Moreover, the Teva Defendants need to review any documents to be produced for privilege and confidentiality, including making any redactions permitted under the Stipulated Electronic Discovery Protocol [Dkt. 127]. Other of Mr. Jaffe's recommendations similarly make no sense.

24. I find it astonishing that Mr. Jaffe could say that "[t]he TEVA Defendants are pioneering CAL review systems within this MDL," and that "Defendants are beta testing this system" (Decl. at 18), when this technology has been in use for a decade—including in many MDLs—and when CAL has been empirically shown to be state of the art and superior to other TAR methods. *See* Gordon V. Cormack & Maura R. Grossman, *Evaluation of Machine-Learning Protocols for Technology-Assisted Review in Electronic Discovery*, in Proc. of the 37th Int'l ACM SIGIR Conference on Rsch. & Dev. in Info. Retrieval 253 (2014). While TAR 2.0 or CAL may be unfamiliar to Plaintiffs and their consultant, it does not render the technology new, novel, or untested in the field of electronic discovery and to those of us who regularly practice in this field.

25. I declare under penalty of perjury under the laws of the United States of America that the foregoing is true and correct.

Executed this 28th day of July, 2020 in Waterloo, Ontario, Canada.

/s/ Maura R. Grossman

Maura R. Grossman, J.D., Ph.D.